Konrad-Zuse-Zentrum
für Informationstechnik Berlin

Takustraße 7
D-14195 Berlin-Dahlem
Germany

MICHAEL MEYER*, THOMAS STEINKE*,
MARIA BRANDL†, JÜRGEN SÜHNEL†

# Density functional study of guanine and uracil quartets and of guanine quartet metal/ion complexes

* Konrad-Zuse-Zentrum
† Institut für Molekulare Biotechnologie

ZIB-Report 00-16 (May 2000)

# Density functional study of guanine and uracil quartets and of guanine quartet metal/ion complexes

Michael Meyer*, Thomas Steinke*
Maria Brandl†, Jürgen Sühnel†

**Abstract**

The structures and interaction energies of guanine and uracil quartets have been determined by B3LYP hybrid density functional calculations. The total interaction energy $\Delta E^T$ of the $C_{4h}$-symmetric guanine quartet consisting of Hoogsteen type base pairs with two hydrogen bonds between two neighbour bases is -66.07 kcal/mol at the highest level. The uracil quartet with C6-H6...O4 interactions between the individual bases has only a small interaction energy of -20.92 kcal/mol and the interaction energy of -24.63 kcal/mol for the alternative structure with N3-H3...O4 hydrogen bonds is only slightly more negative. Cooperative effects contribute between 10 and 25 % to all interaction energies. Complexes of metal ions with G-quartets can be classified into different structure types. The one with $Ca^{2+}$ in the central cavity adopts a $C_{4h}$-symmetric structure with coplanar bases, whereas the energies of the planar and non-planar $Na^+$ complexes are almost identical. The small ions $Li^+$, $Be^{2+}$, $Cu^+$ and $Zn^{2+}$ prefer a non-planar $S_4$-symmetric structure. The lack of coplanarity prevents probably a stacking of these base quartets. The central cavity is too small for $K^+$ ions and therefore this ion favours in contrast to all other investigated ions a $C_4$-symmetric complex, which is 4.73 kcal/mol more stable than the $C_{4h}$-symmetric one. The distance 1.665 Å between $K^+$ and the root mean squares plane of the guanine bases is approximately half of the distance between two stacked G-quartets. The total interaction energy of alkaline earth ion complexes exceeds the ones with alkali ions. Within both groups of ions the interaction energy decreases with an increasing row position in the periodic table. The B3LYP and BLYP methods lead to similar structures and energies. Both methods are suitable for hydrogen-bonded biological systems. Compared with the before mentioned methods the HCTH functional leads to longer hydrogen bonds and different relative energies for two U-quartets. Finally we calculated also structures and relative energies with the MMFF94 forcefield. Contrary to all DFT methods, MMFF94 predicts bifurcated C-H...O contacts in the uracil quartet. In the G-quartet the MMFF94 hydrogen bond distances N2-H22...N7 are shorter than the DFT distances, whereas the N1-H1...O6 distances are longer.

*Konrad-Zuse-Zentrum für Informationstechnik Berlin, Takustr. 7, D-14195 Berlin
†Institut für Molekulare Biotechnologie, Beutenbergstr. 11, D-07745 Jena

## Introduction

Base pairs linked by hydrogen bonds (H-bonds) constitute an important structural motif of RNA and DNA. Therefore, extensive theoretical work has been done to supplement experimental studies. For quantum chemical studies of base pairs Hartree-Fock (HF) theory and Møller-Plesset perturbation theory of second order (MP2) have been used frequently, whereas for individual bases higher theoretical levels have been adopted [1]. As the computational demand increases rapidly with the number of bases, density functional theory (DFT) appears to be an attractive alternative for nucleic acid model systems. The application of DFT methods has provided promising results for individual fragments of organic and bimolecular structures [2, 3, 4, 5, 6]. To our knowledge Santamaria and Vazquez [7] were the first who have applied DFT to the Watson-Crick base pairs guanine-cytosine (GC) and adenine-thymine (AT) using the Vosko, Wilk and Nusair [8] local and Becke Perdew [9, 10] non-local correlation functionals and compared the results with conventional ab initio methods. Then, Šponer and coworkers [11] found for a series of base pairs that the interaction energies derived from the B3LYP hybrid density functional method [12, 13] are much the same as those from MP2 single point calculations when the same geometry is used. We have investigated the H-bonded AT-base pair and the unusual adenine-difluorotoluene (AF) pair and found close a correspondence between B3LYP and single point MP2 base pair interaction energies [14]. The distances between both bases in AT-pairs from B3LYP calculations are in better agreement with experimental AT-pair distances in DNA structures than the ones derived from HF calculations. When compared with a MP2(FC) optimized structure, B3LYP performed better than HF for the geometry of a water-mediated base pair. The corresponding B3LYP interaction energies are only slightly larger than the MP2(FC) data [15]. Similarly, a good agreement between B3LYP and MP2 calculations has been found for the complex of metal ions with the thioguanine-cytosine base pair [16]. These findings encouraged us to apply the B3LYP method to base quartets. To our knowledge this is the first nonempirical quantum-chemical study on base quartet structures.

Guanine (G)-quartets are known to be formed by G-rich telomeric DNA located at the end of eukaryontic chromosomes [17, 18] and by the self-association of guanosine gels [19]. G-quadruplex based inhibitors of telomerase may be relevant to cancer therapy [20, 21]. In addition, G-rich repeats do also occur in other parts of the human genome. Therefore, quadruplex formation may also be important for other diseases than cancer [22]. In this context it is important to note that $NH_4^+$ and metal cations play an important role in the folding and stability of quadruplex structures. Uracil(U)-quartets have not attracted the same interest thus far. An example has been found in an unusually stable RNA tetraplex formed by parallel strands of $r(UG_4U)$ [23]. This tetraplex structure has been determined by two-dimensional NMR spectroscopy and restrained molecular dynamics simulations. It has a fourfold symmetry axis and a practically planar Hoogsteen type geometry of the guanine bases with two H-bonds N2-H22...N7 and N1-H1...O4 between the G-monomers (Fig. 1a). The U-quartet at the 3'-terminus has a non-planar conformation and forms classical N3-H3...O4 H-bonds between the bases, but the U-quartet at the 5'-terminus is planar and linked only by C6-H6...O4 contacts in a cyclic manner (Fig. 2). Such C-H...O interactions are well-known for small organic molecules [24, 25, 26] and have recently attracted much interest in structural biology as well [27, 28]. Often, however, C-H...O contacts are accompanied by strong neighbouring interactions like H-bonds so that the C-H...O contacts may be enforced by the neighbour interaction and do not contribute to the stability of the system on their own. Therefore, the U-quartet found in the RNA tetraplex structure is of particular interest because the bases are connected by C6-H6...O4 contacts only.

We want to supplement previous theoretical studies of C-H...O interactions [29, 30] in a rigorous way and provide also additional data on cooperativity of H-bonded biological systems [31, 32].We have performed B3LYP calculations to determine the structures and interaction energies of the individual bases in the G-quartet and in both U-quartets. As guanine is intrinsically non-planar G-quartets may show small deviations from the coplanar $C_{4h}$-symmetric structures. Therefore, we have also carried out calculations for structures with pyramidal amino groups ($C_4$- and $S_4$-symmetry). For these lower symmetries the G-quartets may show deviations of the base planes from the coplanar geometry. Finally, we have investigated the structural and energetical features of the G-quartet interaction with metal ions in order to analyse the ion dependence of the self-association. The methodical part of our work comprises a comparison of different basis sets and of the performance of the B3LYP density functional method with the BLYP and HCTH approaches based on Becke's 1988 exchange functional [9] and the functional of Hamprecht and coworkers [33], respectively. The latter approaches might provide access to even larger biological systems as a consequence of the computational efficiency method arising from the missing HF step present in B3LYP. The applicability of these methods requires that H-bonded systems can be calculated accurately as H-bonds are of utmost importance in biology [34, 35]. Even medium-sized biopolymers exceed the range of applicability for DFT calculations. We have, therefore, compared the computed structures with the results obtained with the MMFF94 forcefield derived from quantum-chemical calculations ranging between the HF and MP4SDQ single point levels.

# Methods

An initial structure of the quartets has been generated from the coordinates of the RNA tetraplex with the Protein Data Bank code 1rau [23]. The bases have been capped with hydrogen atoms and minor deviations of the G- and U-quartet geometries from $C_{4h}$-symmetry have been removed with UNICHEM [35] (Figs. 1a, 2a). The U-quartet with N-H...O H-bonds (Fig. 2b) has also been investigated at $C_{4h}$-symmetry. The $C_4$- and $S_4$-symmetric G-quartets have been studied for comparison with the coplanar complex structure. The initial structures used for optimizations consist of four G-monomers with a $C_{4h}$-symmetric complex geometry except for pyramidal amino groups. The $C_4$-symmetric quartet has all amino hydrogen atoms on the same side of the base planes whereas the $S_4$-symmetric has hydrogen atoms above and below the base plane in an alternating sequence. The metal ions $Li^+$, $Na^+$, $K^+$, $Be^{2+}$, $Mg^{2+}$, $Ca^{2+}$, $Cu^+$ and $Zn^{2+}$ have been positioned at the centre of the G-quartets for $C_{4h}$- and $S_4$-symmetry and at a distance of 1.6 Å below the center for $C_{4h}$-symmetry prior to the geometry optimizations (Fig. 3).

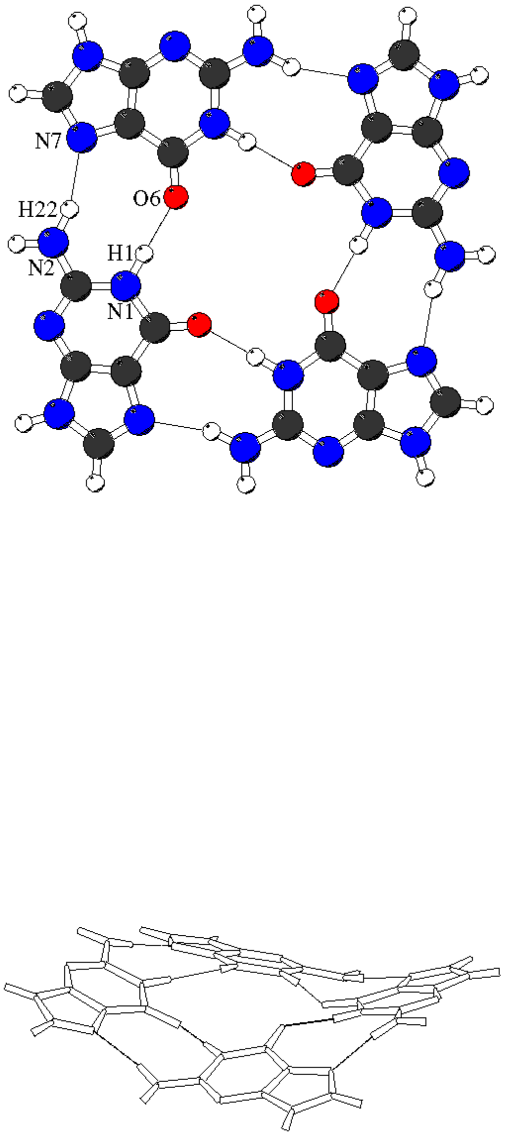


Figure 1: Optimized structure of the guanine quartet with a $C_{4h}$- (a) and $S_4$-symmetry (b). The figures have been created with MOLSCRIPT [59].

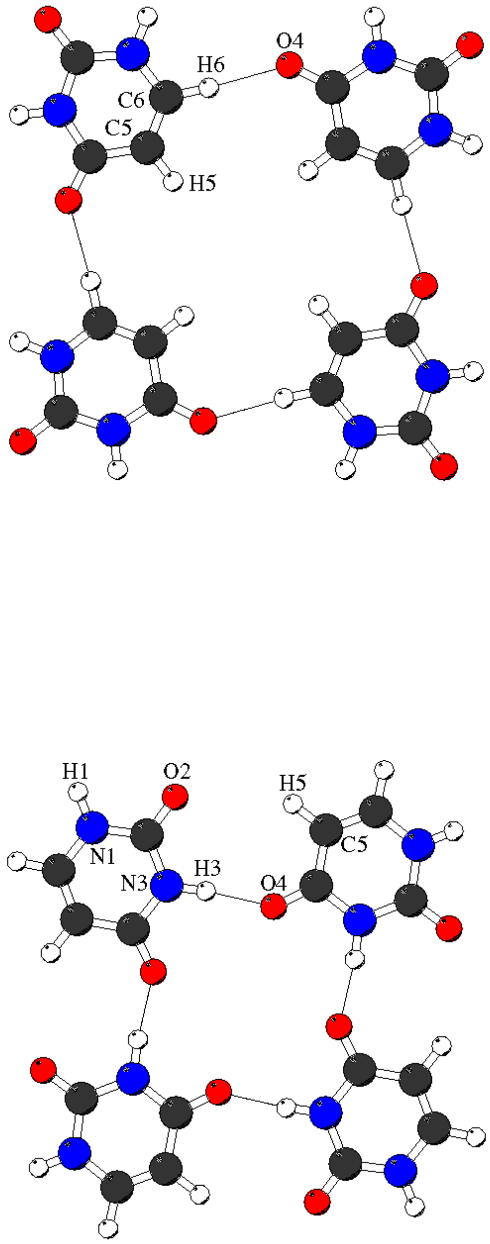


Figure 2: Optimized structure of the uracil quartet with C-H...O interactions (a) and N-H...O hydrogen bonds (b).

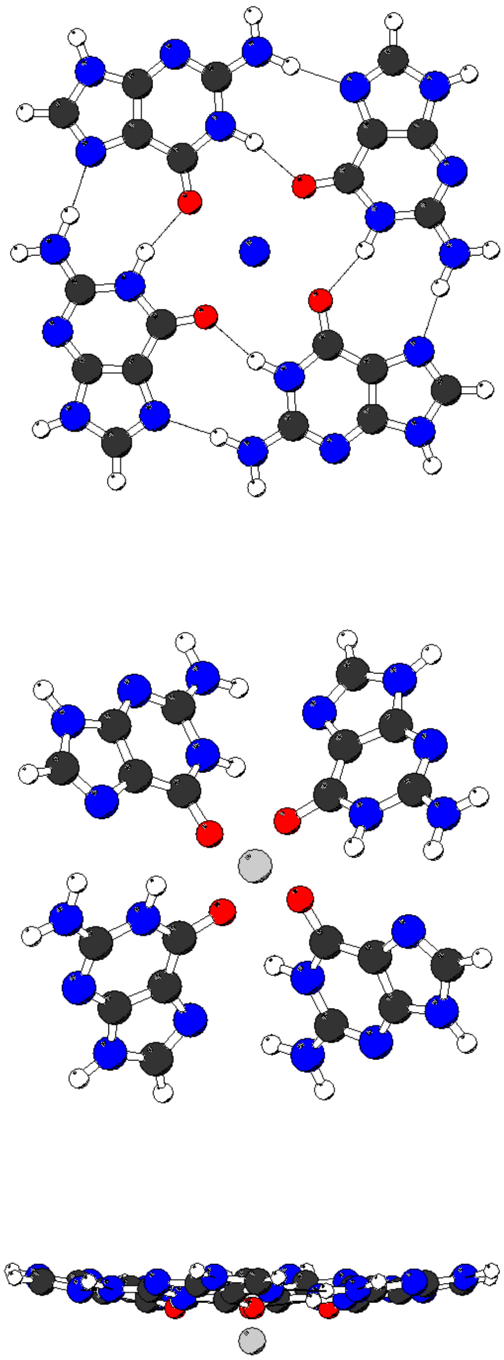

Figure 3: Optimized structures of G-quartets with ions. a) complex with $Na^+$ and $C_{4h}$-symmetry, b) complex with $Be^{2+}$ and $S_4$-symmetry c) complex with $K^+$ and $C_4$-symmetry

The structures of G, U and of the quartet structures have been optimized using the B3LYP hybrid density functional method [12, 13] and the 6-31G(d,p), 6-311G(d,p) and 6-311+G(d,p) basis sets [37, 38, 39, 40]. Individual bases have also been investigated with the MP2(FC) method. Energy minima have been verified with subsequent frequency calculations and interaction energies have been corrected for the basis set superposition error (BSSE) with the standard counterpoise method [41]. Additional DFT calculations were carried out with the BLYP [9, 13] and HCTH [33] functionals and the DZVP and TZVP basis sets optimized for DFT calculations [42]. For the metal ions average relativistic potentials with a large orbital basis and a small core have been used [43, 44]. G-quartet metal complexes with $Li^+$, $Na^+$, $Be^{2+}$ and $Mg^{2+}$ have also been investigated using the all electron approach (AE) with the standard B3LYP/6-31(d,p) basis set to analyse the suitability of effective core potentials. With ECP we designate those calculations with effective core potentials and a DZ basis at the metal ions. A 6-31G(d,p) basis was used throughout for the base atoms in complexes with metal cations. All calculations were carried out with GAUSSIAN94 [45], except for the HTCH-calculations, which were performed with the beta release of DGauss 5.0 [46].

The interaction energy $\Delta$E of the quartets was calculated according to the BSSE correction scheme, Eq. (1), where $\Delta E(B_4)$ denotes the energy of the quartet consisting of four identical bases B and E(B) denotes the energy of a single base computed with the full quartet-centered basis set.

$$\Delta E = E(B_4) - 4E(B) \tag{1}$$

The total interaction energy can also be expressed in terms of pairwise interaction energies.

$$\Delta E = 4\Delta E^n + 2\Delta E^d + \Delta E^c \tag{2}$$

The first two terms describe pair interactions whereas the third term $\Delta E^c$ represents the cooperative contribution. $\Delta E^n$ is the interaction energy between neighbour base pairs, the interaction energy between diagonal opposite base pairs is given by $\Delta E^d$. Each base is deformed slightly from its ideal monomer geometry upon complex formation and the corresponding deformation energy $\Delta E^{def}$ is the energy difference between the structure adopted by a single base in the complex and the optimized structure of this individual base. The total interaction energy $\Delta E^T$ is the sum of $\Delta$E and $\Delta E^{def}$.

$$\Delta E^T = \Delta E + 4\Delta E^{def} \tag{3}$$

Furthermore the zero-point vibrational energy difference $\Delta$ZPE between the quartet and four individual bases contributes to $\Delta E_0$ defined as

$$\Delta E_0 = \Delta E^T + \Delta ZPE \tag{4}$$

The total interaction energy $\Delta E^T$ can be generalized for the quartets with a metal ion M.

$$\Delta E^T = E(MB_4) - 4E(B) - E(M) + 4\Delta E^{def} \tag{5}$$

**TABLE I.** Selected distances (Å) and angles ( $^\circ$ ) of the U- and G-quartets.

| | HCTH/ TZVP | | BLYP/ TZVP | | B3LYP/ TZVP | | B3LYP/ 6-311G(d,p) | | B3LYP/ 6-311+G(d,p) | |
|---|---|---|---|---|---|---|---|---|---|---|
| *U-quartet CHO* | | | | | | | | | | |
| r(H6...O4) | 2.270 | | 2.156 | | 2.115 | | 2.100 | | 2.111 | |
| r(H5...O4) | 4.332 | | 4.213 | | 4.153 | | 4.194 | | 4.121 | |
| $\alpha$(C6-H6...O4) | 172.6 | | 174.7 | | 175.2 | | 172.3 | | 176.7 | |
| $\alpha$(H6...O4=C4) | 133.0 | | 135.3 | | 135.9 | | 132.6 | | 137.5 | |
| *U-quartet NHO* | | | | | | | | | | |
| r(H3...O4) | 1.970 | | 1.804 | | 1.780 | | 1.777 | | 1.795 | |
| r(H5...O2) | 3.179 | | 3.012 | | 2.972 | | 2.955 | | 2.989 | |
| r(O4...O4d) | 5.210 | | 5.024 | | 4.996 | | 4.944 | | 5.041 | |
| $\alpha$(N3-H3...O4) | 177.9 | | 176.6 | | 177.2 | | 176.3 | | 178.3 | |
| $\alpha$(C5-H5...O2) | 144.3 | | 145.2 | | 144.7 | | 145.4 | | 143.9 | |
| *G-quartet* | $C_{4h}$ | $S_4$ | $C_{4h}$ | $S_4$ | $C_{4h}$ | $S_4$ | $C_{4h}$ | $S_4$ | $C_{4h}$ | $S_4$ |
| r(H22...N7) | | 2.325 | 2.054 | 2.089 | 2.109 | 2.123 | 1.971 | 1.992 | 2.015 | 2.039 |
| r(H1...O6) | | 1.941 | 1.815 | 1.795 | 1.816 | 1.798 | 1.782 | 1.755 | 1.791 | 1.771 |
| r(O6...O6d) | | 5.340 | 5.163 | 5.075 | 5.288 | 5.190 | 4.940 | 4.846 | 5.062 | 4.993 |
| $\alpha$(N2-H22...N7) | | 161.1 | 162.9 | 165.4 | 157.4 | 160.7 | 165.8 | 169.1 | 163.2 | 165.8 |
| $\alpha$(N1-H1...O6) | | 173.5 | 178.8 | 178.7 | 173.8 | 175.7 | 176.5 | 174.1 | 180.0 | 176.0 |
| Plane RMS$^{a)}$ | 0.0 | 0.631 | 0.0 | 0.499 | 0.0 | 0.431 | 0.0 | 0.567 | 0.0 | 0.519 |

a) root mean square deviation of all atoms from a least squares plane

Forcefield calculations have been carried out with MMFF94 [47] as implemented in Sybyl 6.4 [48]. A dielectric constant of 1.0 has been used throughout and the optimizations have been terminated at a gradient of 0.001 kcal/mol.

# Results

The calculated geometrical parameters and energies of the monomers and of the metal-free G- and U-quartets are summarized in Tab. I-III. We discuss first the calculated quartet properties obtained with the B3LYP method and Pople type basis sets, because the 6-31G(d,p) basis has been used frequently for base pairs [1, 11, 14, 31, 32]. In a subsequent section we compare the results obtained with B3LYP and different basis sets and then we discuss the data obtained with the different DFT methods and a common basis set.

## U-quartet

The U-quartet has been optimized for two different complex structures at $C_{4h}$-symmetry. One structure has C6-H6...O4 links (CHO structure) between the bases (Fig. 2a) and the other one is linked by N3-H3...O4 H-bonds (NHO structure) and by additional relatively long C5-H5...O2 interactions (Fig. 2b). As a consequence of the weak C6-H6...O4 interactions between the individual bases, three vibrations corresponding to inter base motions with frequencies below 10 cm$^{-1}$ exist at the B3LYP/6-311+G(d,p) level. Yet, the CHO U-quartet has a local minimum for the planar structure shown in Fig. 2a. The H6...O4 distances of 2.100 Å determined with the 6-311G(d,p) basis set are clearly longer than the classical H-bonds distances N3-H3...O4 in the alternative NHO structure (Fig.

**TABLE II.** Total energy E (H), zero point vibration energy ZPE (H), deformation energy $\Delta E^{def}$ (kcal/mol), interaction energy $\Delta E$ (kcal/mol), total interaction energy $\Delta E^T$ (kcal/mol), pairwise interaction energies $\Delta E^n$ and $\Delta E^d$ (kcal/mol), cooperative energy $\Delta E^c$ (kcal/mol) and $E_0$ (kcal/mol) of uracil and the relative energy of uracil quartets with CHO and NHO interactions.

| | HCTH/ TZVP | BLYP/ TZVP | B3LYP/ TZVP | B3LYP/ 6-311G(d,p) | B3LYP/ 6-311+G(d,p) |
|---|---|---|---|---|---|
| *U* | | | | | |
| E | -414.80861 | -414.86528 | -414.95545 | -414.93461 | -414.94607 |
| ZPE | 0.08518 | 0.08347 | 0.08658 | 0.08700 | 0.08669 |
| *U-quartet CHO* | | | | | |
| E | -1659.25866 | -1659.49208 | -1659.85656 | -1659.77799 | -1659.81914 |
| $\Delta E^n$ | | -3.27 | -4.39 | -4.22 | -4.37 |
| $\Delta E^d$ | | -0.44 | -0.48 | -0.47 | -0.48 |
| $\Delta E^c$ | | -3.27 | -3.39 | -3.53 | -3.41 |
| $\Delta E$ | | -19.35 | -21.91 | -21.35 | -21.88 |
| $\Delta E^{def}$ | 0.16 | 0.21 | 0.24 | 0.29 | 0.24 |
| $\Delta E^T$ | | -18.51 | -20.95 | -20.19 | -20.92 |
| $\Delta$ZPE | | 2.06 | 2.08 | 2.05 | 1.97 |
| $\Delta E_0$ | | -16.45 | -18.87 | -18.14 | -18.95 |
| *U-quartet NHO ** | | | | | |
| E | -1659.25925 | -1659.49884 | -1659.86430 | -1659.79008 | -1659.82667 |
| $\Delta E^n$ | | -5.30 | -6.06 | -5.81 | -6.00 |
| $\Delta E^d$ | | -0.02 | -0.02 | -0.08 | -0.03 |
| $\Delta E^c$ | | -2.78 | -2.79 | -2.58 | -2.49 |
| $\Delta E$ | | -24.02 | -27.07 | -25.98 | -26.55 |
| $\Delta E^{def}$ | 0.24 | 0.56 | 0.53 | 0.58 | 0.48 |
| $\Delta E^T$ | | -21.78 | -24.95 | -23.66 | -24.63 |
| $E_{CHO}$-$E_{NHO}$ | 0.37 | 4.24 | 4.86 | 7.59 | 4.73 |

*) no local energy minimum

2b). Nevertheless, these distance are considerably shorter than the sum of van der Waals radii for H and O (2.7 Å). Upon complex formation the C6-H6 bond length increases from 1.075 Å in the monomeric uracil to 1.086 Å in the quartet structure. In the NHO structure there are also rather long C5-H5...O2 distances of 2.989 Å.

Frequency calculations indicate that the U-quartet with C6-H6...O4 contacts corresponds indeed to a local energy minimum, whereas four small imaginary frequencies exist for the complex with N3-H3...O4 H-bonds. The four almost degenerate C6-H6 stretching vibrations are shifted by about 40 $cm^{-1}$ to lower wavenumbers upon complex formation. The total interaction energy $\Delta E^T$ of -20.92 kcal/mol for the U-quartet with C6-H6...O4 contacts is only slightly smaller than the one for the alternative orientation (Fig. 1a), but it is much lower than the one of the G-quartet. As compared to the C6-H6...O4 interactions in the CHO structure the NHO U-quartet has much shorter distances of 1.795 Å between the hydrogen and the acceptor atom. Similar to the O4 atoms in the G-quartet, the distance between both O4 atoms of opposite bases in the U-quartet is 5.041 Å. This means that there is sufficient space for metal ion binding. $\Delta E^T$ amounts to -24.63 kcal/mol. In the NHO U-quartet the pairwise interaction energy between neighbour bases is stronger but the interac-

**TABLE III.** Total energy E (H), zero point vibration energy ZPE (H), deformation energy $\Delta E^{def}$ (kcal/mol), interaction energy $\Delta E$ (kcal/mol), total interaction energy $\Delta E^{T}$ (kcal/mol), pairwise interaction energies $\Delta E^{n}$ and $\Delta E^{d}$ (kcal/mol), cooperative energy $\Delta E^{c}$ (kcal/mol) and $E_0$ (kcal/mol) of guanine and the relative energy of guanine quartets.

| | HCTH/ TZVP | BLYP/ TZVP | B3LYP/ TZVP | B3LYP/ 6-311G(d,p) | B3LYP/ 6-311+G(d,p) |
|---|---|---|---|---|---|
| *G* | | | | | |
| E | -542.44840 | -542.49878 | -542.62995 | -542.56124 | -542.69411 |
| ZPE | 0.11514 | 0.11205 | 0.11654 | 0.11603 | 0.11571 |
| *G-quartet* $C_{4h}$ * | | | | | |
| E | | -2170.09978 | -2170.63602 | -2170.38820 | -2170.91356 |
| $\Delta E^{n}$ | | -12.58 | -13.79 | -14.14 | -13.74 |
| $\Delta E^{d}$ | | -1.93 | -2.25 | -2.01 | -2.10 |
| $\Delta E^{c}$ | | -16.89 | -15.30 | -17.92 | -16.49 |
| $\Delta E$ | | -71.07 | -74.96 | -78.51 | -75.63 |
| $\Delta E^{def}$ | | 2.89 | 2.67 | 2.51 | 2.39 |
| $\Delta E^{T}$ | | -59.51 | -64.28 | -68.47 | -66.07 |
| *G-quartet* $S_4$ | | | | | |
| E | -2169.87558 | -2170.1006 | -2170.63641 | -2170.38882 | -2170.91404 |
| ZPE | | 0.45367 | 0.47140 | | |
| $\Delta E^{n}$ | | -12.17 | -13.42 | -13.86 | -13.59 |
| $\Delta E^{d}$ | | -1.76 | -2.06 | -1.93 | -1.92 |
| $\Delta E^{c}$ | | -16.78 | -15.37 | -17.97 | -16.60 |
| $\Delta E$ | | -68.78 | -73.17 | -77.27 | -74.80 |
| $\Delta E^{def}$ | 1.33 | 2.13 | 2.15 | 2.20 | 2.15 |
| $\Delta E^{T}$ | | -60.23 | -64.57 | -68.47 | -66.20 |
| $\Delta$ZPE | | 3.43 | 3.29 | | |
| $\Delta E_0$ | | -56.83 | -61.28 | | |
| $E_{C4h}$-$E_{S4}$ | | 0.52 | 0.24 | 0.38 | 0.30 |

*) no local energy minimum

tion energy between the oppositely located bases and the cooperative energy is weaker as in the CHO structure (Table II). Additional diffuse functions do not change the calculated interaction energies significantly and therefore they may be omitted for larger nucleic acid complexes like the G-quartet. But they are important for an accurate determination of the relative energy $E_{CHO}$-$E_{NHO}$ of both U-quartets. At the B3LYP/6- 311G+G(d,p) level the NHO U-quartet is 4.73 kcal/mol more stable than the CHO structure. Nevertheless, the more stable structure does not correspond to a local energy minimum as four imaginary frequencies exist for this structure (see discussion section below).

## G-quartet

The G-quartet has been optimized for the Hoogsteen type geometry with classical H-bonds at $C_{4h}$- , $C_4$- and at $S_4$-symmetry (Fig. 1). As all $C_4$-symmetric structures converged to the $C_{4h}$-symmetry, even the one which is stable in the presence of $K^+$, we will not consider it in the further discussion. As a rule the non-planar $S_4$ structure is somewhat more stable than the planar $C_{4h}$ structure. The energy difference is -0.39 kcal/mol at the B3LYP/6-31G(d,p) and -0.30

kcal/mol at the B3LYP/6-311G(d,p) level. At the latter level, the H1...O6 and H22...N7 distances are 1.791 and 2.015 Å, respectively, in the $C_{4h}$-symmetric structure (Tab. I). At $S_4$-symmetry, the H1...O6 distance is shortened to 1.771 whereas the H22...N7 distance is elongated to 2.039 Å. It appears therefore, that the relative strengths of the different H-bonds to N7 and O6 might be one factor determining the relative energy of the planar and non-planar structures. The preference to a non-planar conformation arises also from the fact that a G monomer with a pyramidal amino group is about 0.56 kcal/mol more stable than the planar molecule at the B3LYP/6-311G(d,p) level. MP2(FC)/6-31G(d,p) predicts even an energy difference of 1.47 kcal/mol. The angle between the planes of the three amino group atoms and the one of all atoms except the two amino hydrogens is 36.2 $^\circ$ at the DFT level and 44.0 $^\circ$ at MP2 level. The plane angle is reduced to 26.2 $^\circ$ upon hydrogen bonding in the $S_4$-symmetric complex. The interaction energies shown in Table III differ by approximately one percent between both quartet conformations. For the $C_{4h}$-symmetric structure B3LYP/6-311G(d,p) predicts a total interaction energy $\Delta E^T$ of -66.07 kcal/mol. According to Eq. (2), the interaction energy $\Delta E^n$ of -13.74 kcal/mol between the H-bonded neighbour bases is much stronger than the interaction energy $\Delta E^d$ of -2.10 kcal/mol between opposite pairs without direct H-bonds. The cooperative energy contribution $\Delta E^c$ is -16.49 kcal/mol. Each base is deformed from its ideal monomer geometry upon complex formation and the corresponding deformation energy $\Delta E^{def}$ amounts to 2.39 kcal/mol for a single G. The strong interaction within the G-quartet relative to the U-quartets causes a higher deformation energy.

## G-quartet with metal ions

Relative energies for the G-quartet metal ion complexes are listed in Tab. IV and structures are shown in Fig. 3. Complexes with ions having small radii, especially $Li^+$, $Be^{2+}$ and $Mg^{2+}$, are more stable at $S_4$-symmetry, whereas complexes with ions posessing large radii like $K^+$ and $Ca^{2+}$ have no local minimum for the $S_4$-symmetry. For $Na^+$ the $C_{4h}$- and the $S_4$-symmetric complexes have approximately the same energy and also the structures corresponding to both symmetries are similar. This is indicated by the 0.245 root-mean-square (RMS) deviation of all atoms in the $S_4$-symmetric structure from a least-squares-plane, which is significantly smaller than for all other ions with small radii exceeding the value of 0.567 for the metal-free G-quartet. The energy difference between the $C_{4h}$-and the $S_4$-symmetric complexes correlates with the RMS deviation of the atoms from the least-squares-plane (Table IV). In contrast to all other investigated complexes with ions has the $K^+$ G-quartet complex a local minimum for $C_4$-symmetry which is 4.73 kcal/mol more stable than the one for $C_{4h}$-symmetry (Fig. 3c). Other ions located initially at the $K^+$ position move into the G-quartet plane during the optimization of the complex. Selected geometric parameters of the complex structures at $C_{4h}$-symmetry are listed in Table V. Relative to the metal free complex, the H1...O6 distances increase upon metal binding and all H22...N7 distances decrease, except for $K^+$. The preference of ions with small radii for the $S_4$-symmetric structure is a consequence of the metal ion - O6 attraction. In this particular symmetry much smaller distances between the central ion and O6 can be achieved. For $Li^+$, $Be^{2+}$ and $Mg^{2+}$, for example, these distances are 2.036, 1.645 and 1.996 Å at

**TABLE IV.** Energy difference ΔE between the $C_{4h}$- and $S_4$-symmetric conformation (kcal/mol) of the G-quartet with a central metal ion and root mean square (RMS) deviation (Å) of all atoms from the least squares plane in the $S_4$-symmetric structures.

| metal ion | ΔE | | plane RMS | |
|---|---|---|---|---|
| | AE | ECP | AE | ECP |
| - | -0.39 | - | 0.567 | - |
| $Li^+$ | -3.12 | -3.41 | 0.881 | 0.962 |
| $Na^+$ | -0.04 | -0.03 | 0.245 | 0.271 |
| $K^+$ | - | a,b | - | b |
| $Cu^+$ | - | -1.96 | - | 0.726 |
| $Be^{2+}$ | -4.52 | - | 1.502 | - |
| $Mg^{2+}$ | -4.14 | -5.58 | 0.943 | 0.962 |
| $Ca^{2+}$ | - | b) | - | b) |
| $Zn^{2+}$ | - | -1.09 | - | 1.050 |

a) The energy difference between the $C_{4h}$- and $C_4$-symmetric complex is 4.73 kcal/mol.

b) No local minimum exists for the $S_4$-symmetric structure.

$S_4$-symmetry, whereas the corresponding distances are 2.142, 1.798 and 2.043 Å according to the AE calculations at $C_{4h}$-symmetry. Small metal ions attract the base O6 atoms in an extraordinarily strong manner. In the metal-free G-quartet with $C_{4h}$-symmetry the distance O6...O6$^d$ between the diagonal bases 4.939 Å at the B3LYP/631G(d,p) level. In the $Be^{2+}$ complex this distance, corresponding to twice the M-O6 distance listed in Tab. V, is reduced by about 1.4 Å to 3.596 Å. Among the various metal-ion G-quartet complexes the respective O6...O6$^d$ distances of 4.568, 4.598 and 5.099 Å for $Na^+$, $Ca^{2+}$, and $K^+$ resemble the distances in the metal-free quartet most closely. $K^+$ in the center leads to the smallest change of the O6...O6d distance, but this ion is the only one extending the distance relative to the metal-free structure. A comparison of the $C_4$- and the $C_{4h}$-symmetric complexes with $K^+$ indicates that the central cavity is too small for this ion. When the ion is located outside of the cavity (Fig. 3c) the H-bond distances H22...N7 and H1...O$^d$ relax from 2.080 to 1.918 Å and from 1.934 to 1.895 Å, respectively. Similarly, the O6...O6$^d$ distance of 5.099 Å is reduced to 4.850 Å, which is close to the corresponding distance of 4.939 Å in the metal-free quartet structure . The Mulliken population analysis indicates that there is a large charge transfer from the G-quartet to the central metal ion M (Tab. VI). This charge transfer is larger for the alkaline earth metal ions than for the alkali metal ions. It decreases for both groups with increasing ionic radii. According to Tab. VI the total interaction energy $\Delta E^T$ is drastically increased when a metal ion is present in the G-quartet. $\Delta E^T$ is stronger in the presence of twofold charged ions than for the singly charged ions and it decreases for the alkali and alkaline earth ions from $Li^+$ to $K^+$ and from $Be^{2+}$ to $Ca^{2+}$. This decrease of the interaction energy has its counterpart in the decrease of the deformation energies. Those ions having a large interaction energy are able to deform the individual bases more than those with a small interaction energy.

## Basis set dependence of the results

Comparative calculations with different basis sets have been carried out for the quartets without cations. The distances between the monomers in the

**TABLE V.** Nonbonded distances (Å) of the $C_{4h}$-symmetric G-quartet with metal ions.

| metal ion | $r_{ion}^{a)}$ | r(M-O6) | | r(H22...N7) | | r(H1...O6) | |
|---|---|---|---|---|---|---|---|
| | | AE | ECP | AE | ECP | AE | ECP |
| - | - | - | - | 1.971 | - | 1.782 | - |
| $Li^+$ | 0.68 | 2.142 | 2.139 | 1.799 | 1.796 | 1.855 | 1.852 |
| $Na^+$ | 0.97 | 2.291 | 2.284 | 1.891 | 1.880 | 1.974 | 1.961 |
| $K^+$ | 1.33 | - | 2.545 | - | 2.080 | - | 1.934 |
| | | - | $2.643^b$ | - | $1.918^b$ | - | $1.895^b$ |
| $Cu^+$ | 0.96 | - | 2.290 | - | 1.811 | - | 1.826 |
| $Be^{2+}$ | 0.35 | 1.798 | 1.805 | 1.759 | 1.750 | 2.074 | 2.056 |
| $Mg^{2+}$ | 0.66 | 2.043 | 2.026 | 1.804 | 1.799 | 1.978 | 1.982 |
| $Ca^{2+}$ | 0.99 | - | 2.299 | - | 1.956 | - | 2.008 |
| $Zn^{2+}$ | 0.74 | - | 2.029 | - | 1.774 | - | 1.922 |

a) crystallographic ion radius [60].

b) $C_4$-symmetry

**TABLE VI.** Charge of the metal ion of the G-quartet with a central metal ion.

| metal ion | AE | ECP |
|---|---|---|
| $Li^+$ | 0.481 | 0.569 |
| $Na^+$ | 0.520 | 0.863 |
| $K^+$ | - | 0.874 |
| $Cu^+$ | - | 0.650 |
| $Be^{2+}$ | 0.773 | 1.075 |
| $Mg^{2+}$ | 1.037 | 1.699 |
| $Ca^{2+}$ | - | 1.778 |
| $Zn^{2+}$ | - | 1.326 |

**TABLE VII.** Interaction energy $\Delta E^T$ (kcal/mol) and deformation energy $\Delta E^{def}$ (kcal/mol) of the $C_4$-symmetric G-quartet with a central metal ion.

| metal ion | $\Delta E^T$ | | $\Delta E^{def}$ | |
|---|---|---|---|---|
| | AE | ECP | AE | ECP |
| - | -75.6 | - | 1.86 | - |
| $Li^+$ | -209.9 | | 4.26 | 4.25 |
| $Na^+$ | -188.9 | -185.8 | 3.87 | 3.39 |
| $K^+$ | - | -149.9 | - | 3.46 |
| | - | $-154.6^a$ | - | $3.49^a$ |
| $Cu^+$ | - | -231.8 | - | 4.47 |
| $Be^{2+}$ | -558.6 | | 8.45 | 8.48 |
| $Mg^{2+}$ | -454.4 | -441.2 | 7.18 | 6.87 |
| $Ca^{2+}$ | - | -344.6 | - | 5.78 |
| $Zn^{2+}$ | - | -483.9 | - | 7.77 |

a) $C_4$-symmetry

U-quartets obtained with TZVP, 6-311G(d,p) and 6-311+G(d,p) are in close agreement. Often the TZVP data are in between the results obtained with both Pople type basis sets (Tab. I). For the G-quartet the distances between the individual bases increase when 6-311G(d,p) is used instead of 6-31G(d,p), DZVP gives even larger distances (Tab. II). The U-quartet with C6-H6...O4 interactions shown in Fig. 2a has to be optimized at least with the 6-311G(d,p) basis or with a larger set supplemented by diffuse functions. Smaller basis sets like 6-31G(d,p) are not suitable for this complex with C-H...O interactions, because the optimized structure has imaginary vibrational frequencies at this level and therefore it does not correspond to an energy minimum of the potential energy. The relative total energy $E_{CHO}$-$E_{NHO}$ between both U-quartets shows a noticeable basis set dependence (Tab. II). TZVP provides results, that are more similar to 6-311+G(d,p) than to 6-311G(d,p). Pairwise interaction energies, the cooperative energy and the deformation energy of the three basis sets are in close agreement for the U-quartet. For the G-quartet, the interaction energies obtained with 6-31G(d,p) are somewhat larger than the DZVP and 6-311G(d,p) results. The monomer zero point energies obtained with the different basis sets agree within 0.3 kcal/mol for both quartets, differences in the change of the zero point energies $\Delta$ZPE upon complex formation are even smaller.

## Performance of the DFT methods

The BLYP interaction energies listed in Tab. II and III are in general smaller than the B3LYP results, when the same basis is used. The relative difference depends on the molecule type. For the G-quartet the difference is about 3%, for the U-quartets the differences increases to approximately 12%. HCTH interaction energies were not calculated. However, taking into account the correspondence between the interaction energy and the deformation energy of the monomer, one might expect that the HTCH interaction energies are even smaller than the BLYP interaction energies. The H-bond lengths given in Table I increase from B3LYP over BLYP to HCTH. The difference between the latter methods is quite large. HCTH predicts also the largest non-planarity of the G-quartet. The plane angle of 37.9$^\circ$ between the amino group and the heterocyclic ring system of the G-monomer from HTCH is not much smaller than the B3LYP and BLYP results of 40.9$^\circ$ and 42.3$^\circ$, respectively. Zero point vibrational energies of the monomers are in close correspondence for HTCH and B3LYP, only the BLYP energies are somewhat smaller.

## Forcefield calculations

Selected structure parameters from MMFF94 calculations of the G- and the U-quartets are listed in Tab. VIII for comparison with the DFT data given in Tab. I. All H-bonds from the forcefield calculations are somewhat shorter than the ones from the three DFT methods. For the U-quartet with CH...O interactions the structures are significantly different. In contrast to the DFT calculations predicting only a short contact between H6 and O4 as displayed in Fig. 2a, MMFF94 predicts a bifurcated interaction with short distances of 2.684 and 2.843 Å between H6 and O4 and between H5 and O4. Such a bifurcated orientation of the bases in the U-quartet has also been predicted from the combined NMR and MD-structure of the tetraplex from r($UG_4U$)

**TABLE VIII.** Summary of the forcefield calculations for the G- and U-quartets

| | | |
|---|---|---|
| *U-quartet CHO* | | |
| r(H6...O4) | 2.684 | |
| r(H5...O4) | 2.843 | |
| $\alpha$(C6-H6...O4) | 123.7 | |
| $\alpha$(H6...O4=C4) | 159.1 | |
| *U-quartet NHO* | | |
| r(H3...O4) | 1.665 | |
| r(H5...O2) | 2.915 | |
| r(O4...O4d) | 5.219 | |
| $\alpha$(N3-H3...O4) | 172.4 | |
| $\alpha$(C5-H5...O2) | 140.0 | |
| *G-quartet* | | |
| | $C_{4h}$ | $S_4$ |
| r(H22...N7) | 1.676 | 1.627 |
| r(H1...O6) | 1.907 | 2.020 |
| r(O6...O6d) | 5.013 | 4.868 |
| $\alpha$(N2-H22...N7) | 156.1 | 160.0 |
| $\alpha$(N1-H1...O6) | 180.0 | 169.8 |
| Plane RMS | 0.0 | 0.672 |

(see below). The forcefield energy difference of 19.84 kcal/mol between both U-quartets is significantly larger then the results from all DFT calculations. Similarly, the energy difference $E_{C4h}$-$E_{S4}$ between both G-quartet structures of 6.07 kcal/mol exceeds the corresponding DFT results in Table III significantly.

# Discussion

## Guanine

The DFT calculations on G indicate that this molecule has a pyramidal amino group. The planarity of amino groups has been discussed in detail by Halgren in relation to the MMFF94 forcefield [50]. The plane angles between the amino group and the least squares plane through all other atoms including the amino nitrogen is about 36.2$^\circ$ at B3LYP/6-311G(d,p) level, which is in reasonable agreement with the MP2(FC)/6-31G(d,p) estimate of 44.0$^\circ$. The corresponding energy differences between the puckered and the planar structures are 0.56 and 1.47 kcal/mol, respectively. The calculated non-planarity of G is in line with the pyramidal nitrogen atoms in the microwave structures of aniline (42.4$^\circ$) [51] and methylamine (54.53$^\circ$ in the staggered conformation [52]). On the other hand, many planar amino groups have been found in crystal structures. It may be that the average structure arising from the fast inversion in a symmetric potential is observed in the X-ray experiment. Packing effects in crystals also seem to favour planar structures. For example, the structures of biphenyl and flavonoids are non-planar in the gas phase and solution but almost coplanar in the crystal structures [6].

## G- and U-quartets

The computed structures of the metal-free G-quartets cannot be compared reasonably with experimental data, because the central cavity is probably filled with ions or solvent. It is also not possible to draw comparisons with the experiment for the U-quartet with C6-H6...O4 interactions, because the NMR-data have provided no structure information for this part of the tetraplex [23]. Instead, the U-quartet showing a high flexibility at the 5'-terminus of the PDB entry 1rau should be regarded as an MD result. In fact, there are some differences in the structures obtained by both techniques. The MD result is a bifurcated conformation with both hydrogen atoms H5 and H6 involved in H-bonds of with O4 of the neighbour U. The corresponding hydrogen acceptor atom distances of 2.873 and 2.927 Å resemble the bifurcated structure derived from MMFF94 (see above). The DFT calculations, however, indicate that there is only a single H-bond with a short H6...O4-distance of 2.111 Å, whereas the H5...O4-distance is 4.121 Å. But one has to keep in mind that, if the interaction energy is low like in the unusual AF pair, there may be a disagreement between calculated geometry of a fragment in the gas phase [14] and the experimental geometry [54] of the corresponding biopolymer due to interactions with the neglected environment. The U-tetraplex has indeed a low total interaction energy $\Delta E^T$ of about -21.80 kcal/mol and a pairwise interaction energy of -4.36 kcal/mol between neighbour bases. Thus, the U-quartet geometry might be influenced noticeably by the environment. The U-quartet with N3-H3...O4 H-bonds has a distance between the opposite oxygen atoms O4 of the cavity of 5.040 Å. This is comparable to the distances between the corresponding distances of 5.061 Å in the planar G-quartet and 4.993 Å in the $S_4$-symmetric structure. Thus, there is enough space for metal coordination in both quartets. As shown in Fig. 4, the electrostatic potential has minima at the centre of the G-quartet and at the centre of the U-quartet with N3-H3...O4 H-bonds. It is somewhat more negative in the former quartet, which enables a stronger interaction with a central cation.

The total interaction energy of the G-quartet is -66.07 kcal/mol corresponding to -13.74 kcal/mol between each pair in the quartet. This pairwise interaction energy is in the range of the interaction energies between -9.2 and -24.8 kcal/mol for G-pairs, determined previously for three relative orientations different from the one in the tetraplex [55]. The total U-quartet interaction energy -20.92 kcal/mol is much lower, confirming that C-H...O interactions are much weaker than classical H-bonds. For example, the pairwise interaction energy -4.36 kcal/mol between next neighbour bases is only about -2 kcal/mol stronger than the diagonal interaction energy $\Delta E^d$ of -2.10 kcal/mol within a G-pair without any direct H-bonds (Tab. II and III) and it is approximately 9 kcal/mol weaker than the pairwise interaction energy in a G-neighbour pair forming two H-bonds. The weak interaction energy of the U-quartet is in line with a much lower vibrational frequency shift, which is often used as a criterion for hydrogen bonding. The calculated red shift of the C6-H6 stretching vibration is about 40 $cm^{-1}$, for the C6-H6...O4 system. In contrast, the N1-H1 and N3-H3 stretching modes of U are lowered by 224 $cm^{-1}$ and 212 $cm^{-1}$ in U dihydrate [56]. As the more stable U-quartet structure with N3-H3...O4 H-bonds does not correspond to a local energy minimum, no frequency shifts can be given for this complex. Therefore, another U-quartet structure corresponding to the global energy minimum must exist. For U-dimers it has been possible to locate the global minimum structure with two N1-H1...O2 H-bonds between the bases [57]. However this complex structure does not correspond to biopolymers because N1 is linked to the sugar-phosphate backbone in nucleic acids.

Optimizations of the G-quartet starting from $C_4$-symmetry converged to planar $C_{4h}$- symmetric structures. As a consequence of pyramidal amino groups of the monomers, G-quartets with $S_4$-symmetry are the most stable structures. But the energy difference listed in Table III is so small that it can easily be compensated by stacking effects. Furthermore, planar structures can be stabilized by cations.

Common to the investigated quartets is a high contribution of non-additive contributions to the interaction energy $\Delta E^T$, which amounts to 25 % for the G-quartet. 16 and 10 % are the cooperative contributions for the U-quartet. In another cyclic system, the non-planar water-mediated uracil-cytosine base pair consisting of two bases and a water molecule replacing a single direct H-bond between the pair, a cooperative contribution of 13 % has been found [31]. So we can conclude that bases linked by both classical H-bonds and C-H...O interactions can lead to cooperative effects of comparable magnitude. But as we determine a global energy for the interaction of the bases, we cannot attribute this to the H-bonds alone. On the other hand, Šponer and coworkers described H-bonded trimers of DNA bases like T.AT and G.AT showing almost no cooperativity [32]. This raises the intriguing question of the derivation of a model explaining the magnitude of the cooperative contribution in nucleic acids. As a result of the cooperative effects the accuracy of the energies derived from forcefield methods with fixed parameters may be affected by cooperative energy contributions.

## G-quartet with metal ions

Previously the interaction of metal ions with G-quartets has been discussed often in relation to ionic radii. It has been concluded from NMR-spectra [17] that $Na^+$ has the optimal size to fit into the central G-quartet cavity, whereas $K^+$ is believed to be too big. Therefore, it adopts a position between two stacked G-quartet layers and interacts with the oxygen atoms of two G-quartets. On the other hand, the X-ray structure of d($TG_4T$) refined to 0.75 Å resolution indicates clearly, that the $Na^+$ positions in the central channel of the parallel-stranded tetraplex (PDB entry 352D [53]) are sequence dependent. At the termini the $Na^+$ ions occupy the central position within the plane of a single quartet like in the model systems of this study, whereas in the remaining part they are located between two consecutive G-quartets. Finally it has been shown that ions with small radii like $Li^+$ inhibit the self association of guanosine gels [19].

Our calculations might explain the latter effect in a qualitative manner. Ions with small radii prefer the formation of especially non-planar complexes, which prevents a stacking of G. For $Be^{2+}$ these non-planar structures (Fig. 3b) are 4.5 kcal/mol stable than the planar ones. The origin of this preference for non-planar structures is the formation of short contacts between the central cation and the oppositely charged oxygen atoms O6, which is hindered by the voluminous bases in the planar conformation. Only for $Na^+$, the planar an the non-planar structures are of almost equal energy. Furthermore, the $S_4$-symmetric structure shows the smallest deviation of the least squares plane through all atoms. For $K^+$ and $Ca^{2+}$ no structures of $S_4$-symmetry exist. The RMS- fit of all G-quartet atoms in the complexes with $Na^+$, $K^+$ and $Ca^{2+}$ on those of the metal-free G-quartet is less than 0.2 Å for all three complexes. For $Na^+$ the calculated M-O6 distances of 2.291 Å and 2.284 Å from all electron and effective core potential calculations are in reasonable agreement with the average experimental distance of 2.307 (11) Å in the G-quartets of d($TG_4T$) close to the terminal with non-stacked T-residues. The calculated N2/N7 (AE:2.917, ECP:2.906 Å) and N1/O6 distances (AE:2.879, ECP:2.867 Å) are also in good agreement with the corresponding X-ray distances of 2.879 (13) and 2.863 (32) Å. The single standard error of the average experimental distances of is given in brackets in units of the last significant digit. For $K^+$ complex structures of $C_4$ and $C_{4h}$-symmetry may exist. The latter is 4.73 kcal/mol less stable because the ion causes an elongation of the hydrogen bonds between the G-monomers. The presence of $K^+$ outside of the quartet center causes a small deviation of the coplanar base geometry (plane RMS 0.294 Å) which has also been found in crystal structures with $Na^+$, when the ion is located between two stacked quartets. The distance 1.665 Å between the RMS plane of the G-quartet and the $K^+$ ion corresponds approximately to the half distance between two stacked G-quartets. Therefore, the $K^+$ ion can be complexed in an optimal way by eight O6 atoms of two stacked G-quartets. The electrostatic potential at the centre of the U-quartet with N-H...O H-bonds indicates, that this position is also favourable for the electrostatic interaction with cations. The Figs. 4a and 4c show that regions with the most negative electrostatic potential are a consequence of the close vicinity of the four oxygen atoms O4 in the U-quartet with N-H...O H-bonds and in the G-quartet. In the alternative orientation of the U-quartet with C-H...O interactions the oxygen atoms are far apart from each other and the regions with the lowest electrostatic potential are much less negative. The Figures indicate also, that regions with a negative electrostatic potential exist at the centre above and below the planes of a single quartet. Therefore, stacked quartets may provide an especially negative potential between two quartets, where in fact ions have been observed in the central region of the tetraplex formed by d($TG_4T$) [53]. The electrostatic potential explains also the experimental fact, that $Na^+$ ions are located in the G-quartet plane at the termini. For the terminal quartet a stacked neighbour is missing on one side and therefore the electrostatic potential is more negative in the quartet plane than outside. Consequently, the ions can be found in the plane at the termini and between the planes in the centre of the tetraplex.

An inspection of Tab. VI shows that there is a substantial charge transfer from the G-quartet to the metal ions. This seems to be a general feature of G-metal ion complexes, as it has also been observed in GC-pairs with metal ions [58]. Instead of the formal charge of +1 for the alkali ions and +2 for the alkaline earth ions, there is a much lower positive charge on the metal ion in the complex complexes and the bases loose negative charge. A generalization of the energy decomposition Eq. (2) to quartets with metal ions, however, refers to the reference state of a neutral base and a metal cation with the corresponding formal charge instead of non-integer charged species. Therefore, we have not carried out a calculation of the pairwise interaction energies between the metal cation and the G-molecules. Tab. VI shows that the charge transfer increases from the alkali to the alkaline earth metals.

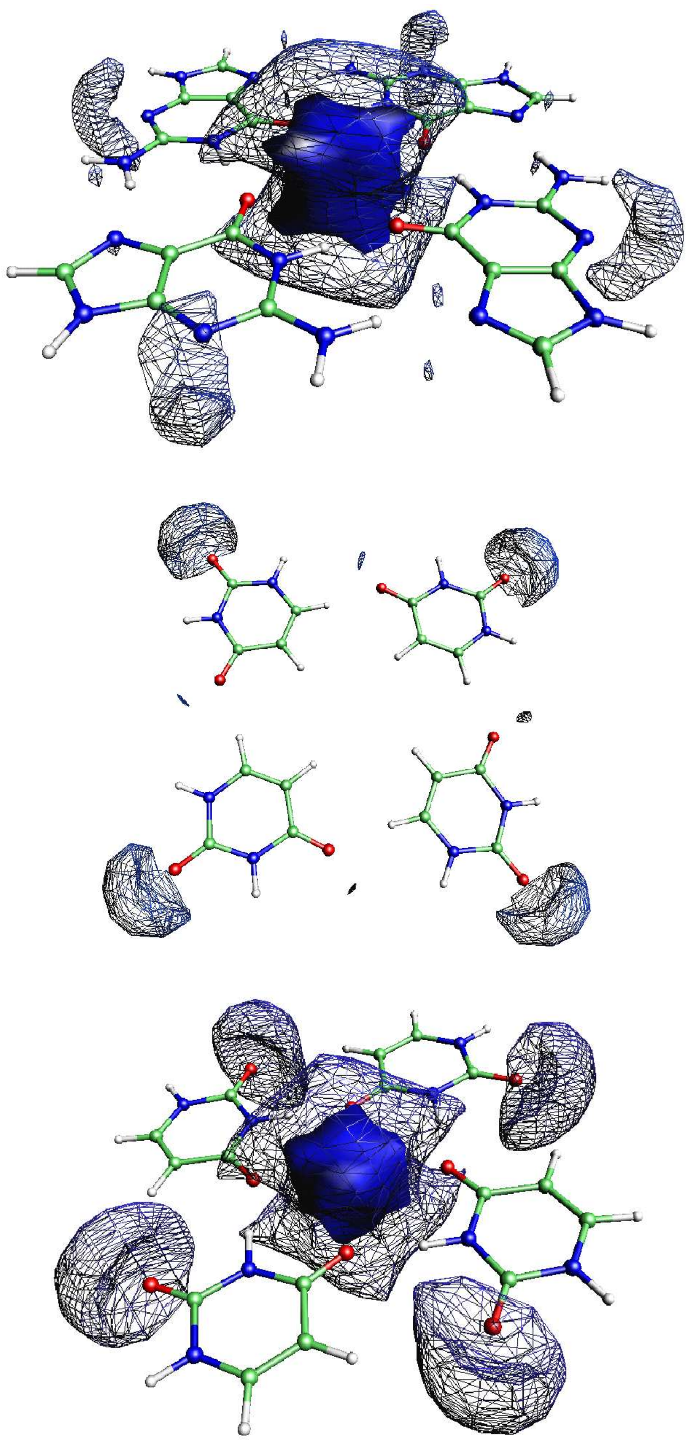

Figure 4: Electrostatic potential at -50 a.u. (opaque surface) and -25 a.u. (wireframe) calculated with B3LYP/6-311G(d,p) for the guanine (a) and the uracil quartets with C6-H6...O4 (b) and N3-H3...O4 (c) interactions.

## Methodological aspects

Conventional ab initio and DFT methods have proved to be useful to get insight in the interaction between nucleic acid bases [34, 49], because they provide data that are not available by experiments, e. g. information about the interaction energies and cooperative effects. Even though the calculations are limited to small fragments without environment, a close agreement between computed and experimental geometrical parameters can be obtained, if the interaction energy is high. The B3LYP density functional method seems to be somewhat more accurate than the HF method [14]. For the basis sets we note that there is a good correspondence between the structures and energies obtained with the DZVP and the larger 6-311G(d,p) basis and between the data derived from TZVP and 6-311+G(d,p). Therefore, these basis sets optimized for DFT calculations provide an efficient and accurate alternative to the more established basis sets. But as the 6-31G(d,p) basis has been widely used by us [6, 14, 15, 31] and others [1, 11, 32, 49] , we have also employed this one for consistency with previous work. A comparison of the results obtained with the AE and the ECP calculations for the G-quartet with metal ions of low atomic number, for which the relativistic effects are probably small, indicates that the average relativistic core potentials [43, 44] are well suited for the calculations of structures and relative energies for bioinorganic compounds (Tab. IV, V). However, the charges from Mulliken population analyses listed in Tab. VI are significantly different. B3LYP has been shown to give results close to MP2 for hydrogen bonded base pairs [14, 15] which is a widely accepted level for the investigation of nucleic acid complexes. BLYP provides somewhat lower interaction energies and somewhat longer H-bonds for the U-quartet, whereas the G-quartet H-bonds are shorter. HTCH calculations lead to much larger distances between the bases of the quartet, whereas the monomer data like the G-plane angle, the zero-point energy and the vibration frequencies are in close agreement with the other DFT methods. Thus, our preliminary analysis indicates that the method appears to be suitable for individual organic molecules, but not for the investigation of complexes. The strengths and the weakness of the HCTH functional should be analysed carefully in further investigations. The forcefield calculations lead to somewhat shorter H-bonds than the three DFT methods and they overestimate the energy difference between both geometries of the G-quartet and the energy difference between both U-quartets, when compared to the DFT methods. For the U-quartet with C6-H6...O4 interactions forcefield calculations predict a different geometry than DFT calculations. To our knowledge experimental reference data are missing in order to decide, which structures are more accurate. As MMFF94 has been parametrized to ab initio structures and energies at a quite high level, it would be interesting in future to carry out quantum-chemical studies on quartet structures with higher level methods and to compare the results with the forcefield calculations. The charge transfer between the bases and the metal ions point out the importance of polarizable charges in forcefields.

# Conclusion

B3LYP and BLYP calculations lead to similar results for the G- and U-quartets. As the former method is also in good agreement with MP2 calculations [6, 14, 16, 31], both methods appear to be suitable for bioorganic and bioinorganic systems with H-bonds and CH...O interactions. DZVP and TZVP basis sets are a noteworthy alternative to the 6-31G(d,p) set, a widely employed basis for the study of nucleic acid bases. Preliminary HCTH calculations yield different results for H-bond lengths and relative energies of complexes. The G-quartet interaction energies are much stronger than the U-quartet energies in both investigated orientations. The same holds for the pairwise components of the interaction energy. Metal cations with a small radii and a high charge enforce a non-planarity of the base quartets and may thus prevent a stacking of G-quartet, unlike $Na^+$ and $K^+$. There is a substantial charge transfer between the G-quartets and the metal ions. The electrostatic potential of G-quartets provides probably favorable binding sites for metal ions between the stacked quartets, wheras isolated quartets have the region of most negative electrostatic potential in the cavity of the quartet centre. U-quartets in the orientation with N3-H3...O4 H-bonds are probably also capable of binding cations at the centre.

# Acknowledgement

We would like to thank Prof. P. B. Moore (Yale University) for a helpful discussion and Oxford Molecular for providing the beta version of DGauss 5. A part of the basis sets was obtained from the Extensible Computational Chemistry Environment Database, as developed and distributed by the Molecular Science Computing Facility, Environmental and Molecular Sciences Laboratory which is part of the Pacific Northwest Laboratory, P. O. Box 999, Richland, Washington 99352, USA.

# References

[1] Šponer, J; Leszcynski, J.; Hobza, P. J Biomol Struct Dyn 1996, 14, 117-135.

[2] St-Amant, A. Density functional methods in biomolecular modeling, in: Reviews in computational chemistry 7 (Ed. K. B. Lipkowitz and D. B. Boyd), VCH, New York 1996, chapter 5.

[3] Berthier, G; Cadioli, B; Gallinella, E.; Amouche, A; Ghomi, M. Theochem 1997, 390, 11-21.

[4] Frisch, M. J.; Trucks, G. W.; Cheeseman, J. R. Systematic model chemistries based on density functional theory: Comparison with traditional models and with experiment, in: Recent developments and applications of modern density functional theory, Theoretical and Computational Chemistry Vol. 4 (Ed. J. M. Seminario), Elsevir, Amsterdam, 1996, p. 679-707.

[5] Meyer, M. Theochem 1997, 417, 163-168.

[6] Meyer, M. Int J Quantum Chem 2000, 76, 724-732.

[7] Santamaria, R. ; Vazquez, A. J Comput Chem 1994, 15, 981-996.

[8] Vosko, S. H.; Wilk, L.; Nusair, M. Can J Phys 1980, 58, 1200-1211.

[9] Becke, A. D. Phys Rev A 1988, 3098-3100.

[10] Perdew, J. P. Phys Rev B 1986, 33, 8822-8824.

[11] Šponer, J; Leszcynski, J.; Hobza, P. J Phys Chem 1996, 100, 1965-1974.

[12] Becke, A. D. J Chem Phys 1993, 98, 5648-5652.

[13] Lee, C.; Yang, G.; Parr, R. G. Phys Rev B 1988, 37, 785-789.

[14] Meyer, M.; Sühnel, J. J Biomol Struct. Dyn 1997, 15, 619-624.

[15] Brandl, M.; Meyer, M.; Sühnel, to be published.

[16] Šponer, J; Burda, J. V.; Leszcynski, J.; Hobza, P. J Biomol Struct Dyn 1999, 17 61-77.

[17] Williamson, J. R. Annu Rev Biophys Biomol Struct 1994, 23, 703-730.

[18] Rhodes, D., Giraldo, R., Curr Opin Struct Biol 1995, 5,311-322.

[19] Guschlbauer, W., Chantot, J.-F.; Thiele, G. J Biomol Struct Dyn, 1990, 8, 491-511.

[20] Read, M. A. ; Wood, A. A.; Harrison, J. R.; Gowan, S. M. ; Kelland, L. R.; Dosanjh, H. S.; Neidle, S. J Med Chem 1999, 42, 4538-4546.

[21] Han, H; Hurley, L. H. Trends Pharm Sci 2000, 21, 136-142.

[22] Gilbert, D. E., Feigon. J Curr Opinion Struct Biol 1999, 9, 305-314.

[23] Cheong, C.; Moore, P. B. Biochemistry 1992, 31, 8406-8414.

[24] R. Taylor, R.; Kennard, O. J Am Chem Soc 1982, 104, 5063-5070.

[25] Steiner, T. Chem Commun 1997, 727-734.

[26] Gu, Y.; Kar, T; Scheiner, S. J Am Chem Soc 1999, 121, 9411-9422.

[27] Wahl, M. C.; Sundaralingam, M. Trends Biochem Sci 1997, 22, 97-102.

[28] Brandl, M.; Meyer, M; Sühnel, J. Theoret Chem Acc 1999, 101 103-113.

[29] Starikov, E.; Steiner, T. Acta Cryst D 1997, 53, 345-347.

[30] Ornstein, R. L.; Zheng, Y.-J. J Biomol Struct Dyn 1997, 14, 657-665.

[31] Brandl, M.; Meyer, M.; Sühnel, J. J Am Chem Soc 1999, 121, 2605-2606.

[32] Šponer, J.; Burda, J. V.; Leszczynski; Hobza, P. J Biomol Struct Dyn 1997, 14, 613-628.

[33] Hamprecht, F. A.; Cohen, A. J.; Tozer, D. J. J Chem Phys 1998, 109, 6264-6271.

[34] Schuster, P; Wolschann, P. Chemical Monthly 1999, 947-960.

[35] Jeffrey, G. A.; Saenger, W.; Hydrogen bonding in Biological Structures, Springer Verlag, Berlin 1991.

[36] UNICHEM, Oxford Molecular 1999.

[37] Hehre, W. J.; Ditchfield, R.; Pople, J. A. J Chem Phys1972, 56, 2257-2261.

[38] Hariharan, P. C.; Pople, J. A. Theoret Chim Acta 1973, 38, 213-222.

[39] Krishnan, R.; Binkley, J. S.; Pople, J. A. J Chem Phys 1980, 72, 650-654.

[40] Clark, T.; Chandrasekhar, J.; Spitznagel, G. W.; v. R. Schleyer, P. J Comput Chem 1983, 4, 294-301.

[41] Boys, S. F.; Bernardi F. Mol Phys 1970, 19, 553-577.

[42] Godbout, N; Salahub, D. R.; Andzelm, J; Wimmer, E. Can. J Chem 1992, 70,560-

[43] Pacios, L. F.; Christiansen, P. A. J. Chem Phys 1985, 82, 2664-2671.

[44] Hurley, M. M.; Pacios, L. F.; Christiansen, P. A.; Roos, R. B.; Ermler, W. C. J Chem Phys 1986, 84, 6840-6853.

[45] Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Gill, P. M. W.; Johnson, B. G.; Robb, M. A.; Cheeseman, J. R.; Keith, T. A.; Petersson, G. A.; Montgomery, J. A.; Raghavachari, K.; Al-Laham, M. A.; Zakrzewski, V. G.; Ortiz, J. V.; Foresman, J. B.; Cioslowski, J.; Stefanov, B. B.; Nanayakkara, A.; Challacombe, M.; Peng, C. Y.; Ayala, P. Y; Chen, W.; Wong, M. W.; Andres, J. L.; Replogle, E. S.; Gomperts, R.; Martin, R. L; Fox, D. J.; Binkley, J. S.; Defrees, D. J.; Baker, J.; Stewart, J. P.; Head-Gordon, M.; Gonzalez, C.; Pople J. A.; GAUSSIAN 94, Revision B.3, Gaussian, Inc., Pittsburgh PA (1995).

[46] DGauss 5.0 beta, Oxford Molecular Group Inc., 1999.

[47] Halgren, T. A. J Comput Chem 1995, 17, 490-519.

[48] Sybyl 6.4, Tripos Associates.

[49] Hobza, P; Šponer, J Chem Rev 1999, 99, 3247-3276.

[50] Halgren, T. A. J Comput Chem 1999, 20, 730-748.

[51] Roussy, G; Nonat, A. J Mol Spectrosc 1986, 118, 180-188.

[52] Kreglewski, M. J Mol Spectrosc 1989, 133, 10-21.

[53] Philips, K.; Orozco, M; Luque, F. Q.; Teat, S. J.; Clegg, W.; Luisi, B; to be published.

[54] Guckian, K. M.; Krugh, T. R.; Kool, E. T. Nat Struct Biol 1998, 5, 954-959.

[55] Šponer, J.; Leszczynski, J; Hobza, P. J Phys Chem 1996, 100, 1965-1974.

[56] Aamouche, A; Berthier, G.; Cadioli, B.; Gallinella, E.; Ghomi, M., Theochem 1998, 426, 307-312.

[57] Kratochvil, M; Engkvist, O.; Šponer, J.; Jungwirth, P.; Hobza, P. J Phys Chem A 1998, 102, 6921-6926.

[58] Burda, J. V; Šponer, J.;. Leszczynski, J. ; Hobza, P. J Phys Chem B 1997, 101, 9670-9677.

[59] Kraulis, P. J Appl Crystallogr 1991, 24, 946-950.

[60] Handbook of Chemistry and Physics, 60th ed. , CRC Press, Boca Raton, 1979, p. F214.